Response to comments by Anonymous Referee #2
(The reviewer's comments are in regular text and our response in italics)

This is an interesting, if challenging, exercise in numerical data transformation. The results are worth reporting, even if the exercise does not appear to have been especially successful.

also final comment: In summary, with modified conclusions, this exercise is worth reporting, but largely because it highlights the difficulties and challenges of using pollen data on their own as a palaeo-fire proxy.

We disagree that this exercise has been unsuccessful. We have shown that there is independent evidence about fire in the pollen records (please see comments below about the interpretation of the CCA analyses) and that it is possible to reconstruct changes in fire regimes from the pollen data. (Please see additional comment below about the comparisons with charcoal data).

Since some of the reviewer's comments suggest a misunderstanding of the approach and how to interpret these results, we will explain this logic more explicitly in the revised manuscript.

What should be removed is the claim in the conclusion/abstract that "this new method opens up the possibility of reconstructing changes in fire regimes quantitatively from pollen records" in regions where charcoal data are lacking. The pollen-burnt area relationship that they established for the Iberian Peninsula is not transferrable to other geographical regions, even to adjacent regions such as France. For example, wildfire in most of Iberia is fuel-limited so that burning increases at times of wetter climate, when biomass increases. In contrast, in most of France, there is abundant plant biomass so that fires are caused by drought conditions when vegetation becomes more flammable. There are some specific plant taxa that are fire-sensitive or fire-tolerant such as Cistus monspeliensis, but these indicator species are the exception not the norm. In reality, what the authors have done is to use pollen and charcoal data in combination to fill spatial gaps in coverage within Iberia. They have not shown that charcoal can be replaced by pollen as a palaeo-fire proxy in other regions (e.g., Greece) where charcoal data are lacking. The analysis carried out for Iberia could, in theory, be scaled up to cover larger regions, but they are not transferable from one part of the planet to another. Sub-Saharan Africa, for example, is deficient in charcoal records although African savannahs account for almost half of all wildfires globally. This deficiency cannot be resolved by using the pollen-burnt area relationship in, for example, North America, and transferring it to African pollen records.

The reviewer notes, correctly, that the limitations on fire differ among regions and that it would not make sense to transfer the same relationships from one region to another. However, we are not advocating this. Instead, our analysis shows that there is sufficient information in the vegetation records to derive information about fire. Therefore, the same methodology could be used in other regions, provided that there are entities with both pollen data and modern charcoal data. The absence of charcoal records in some regions, for example the Sahel, means that this approach cannot currently be used to reconstruct palaeo-burnt area there – but there are other regions of the world where it could be applied.

To avoid confusion, we will revise the text in both the Abstract and the Conclusions, as follows:

(abstract) This new method opens up the possibility of reconstructing changes in fire regimes quantitatively from pollen records, after regional calibration of the vegetation-burnt area relationship, in regions where pollen records are more abundant than charcoal records.

(conclusions) The good predictive power of the fxTWA-PLS derived fire-vegetation relationship and the plausibility of the palaeofire reconstructions with respect to known climate changes in the region suggest that this calibration approach could be applied more generally to provide quantitative reconstructions of past fire regimes in other regions where there are limited charcoal data, and pollen data are more abundant.

There are some specific plant taxa that are fire-sensitive or fire-tolerant such as Cistus monspeliensis, but these indicator species are the exception not the norm.

We disagree with the reviewer. The relative importance of fire-adapted taxa in the vegetation assemblage varies with fire regime (see e.g. Harrison et al., 2021 for an analysis of the abundance of fire-adapted respouting species across Europe in general). The BROT database (Tavsanoglu and Pausas, 2018) provides information about fire-adapted species in the Mediterranean region, including those that are fire-resistant because they have thick bark (e.g. Quercus suber*), resprouters (e.g.* Juniperus oxycedrus*,* Smilax aspera*,* Chamaerops humilis*,* Olea europaea*,* Arbutus unedo*), taxa that require fire because they are serotinous (e.g.* Pinus halepensis*,* Pinus pinaster*), and taxa that are stimulated to germinate by smoke or by heat (e.g.* Cistus albidus*,* Cistus monspeliensis*,* Ulex parviflorus*,* Rosmarinus officinalis*). In regions where there is little fire today, the taxa do not display fire adaptations but can be considered sensitive to fire, so there will be a shift in the vegetation assemblage after fire; indeed this is already in some regions where fire frequency has increased recently.*

The fact that many taxa are fire-adapted makes it possible to derive independent information on fire from the pollen assemblages. We have shown that the variance in pollen assemblages that is explained by fire in Iberia is only 1%, which is substantially less than that attributable to other factors (18%) (climate, vegetation, human activities). This is nonetheless sufficient to be able to exploit the pollen assemblages to reconstruct changes in fire regimes. We expect that in more fire-prone regions, where the abundance of fire-adapted vegetation is greater, the proportion of explained variance would be higher. The fact that it works for Iberia, which as the reviewer points out is not the most fire-prone region of the world, is one reason we suggest our approach could be used effectively elsewhere.

At a more fundamental level, trying to use pollen data as a fire proxy also means that pollen cannot then be used to test vegetation-wildfire dynamics and relationships, in the way that Connor et al (2019) did for Iberia during the Holocene. Pollen data, on their own, are not able to provide both cause and effect without falling into the trap of circular reasoning.

On the contrary: the multivariate nature of pollen assemblages means that is entirely possible to infer several quantities simultaneously from the data. Vegetation responds to multiple aspects of the environment, including seasonal climates, fire and other forms of disturbance, and human activities. This has been explored most extensively with respect to climate. Some taxa are particularly sensitive to winter temperatures, for example, while some taxa occur over a wide range of winter temperatures but are sensitive to plant-available moisture. This differential sensitivity to individual climate variables is what allows us to make reconstructions of multiple aspects of the climate from pollen assemblages, and is illustrated in the GAM-based analyses of the climate space occupied by individual European pollen taxa by Wei et al. (2020), Ecology*. The CCA analyses reported in our paper (Table S1) show that in addition to the climate, vegetation and human influences, the pollen assemblages contain information on fire – thus allowing us to use them to reconstruct fire regimes, without any danger of circularity.*

We did not explicitly comment on the use of pollen data to test wildfire-vegetation dynamics in our original manuscript (except as a motivation for using Iberia in the Introduction, line 75 in the original manuscript), but nevertheless we do not agree with *the reviewer that using the pollen to reconstruct burnt area precludes an analysis of vegetation-wildfire dynamics. Analyses of modern controls on burnt area, including the GLM presented in our manuscript, indicate that gross primary production and the relative abundance of grasses are the most important aspects of the vegetation cover in determining fire occurrence and burnt area (see e.g. Bistinas et al., 2014, Biogeosciences; Forkel et al., 2019, Biogeosciences). Although it has been argued that species composition has an impact on fire regimes in different regions within the same biome (e.g. between North American and Siberian boreal forests), this appears to relate to differences in fire adaptations of individual species rather than being a function of the overall vegetation assemblage. Thus, we argue that it would be useful and interesting to examine wildfire-vegetation dynamics with respect to changing abundance of plant functional types and fire adapted taxa.*

Since the issue of circularity may be something that concerns other readers, we will add a paragraph in the Discussion about this issue, as follows:

We have shown that it is possible to derive trends in burnt area through time by applying a quantitative relationship between pollen assemblages and charcoal-derived burnt area to palaeo-vegetation records from the Iberian Peninsula. Our analyses exploit the multivariate nature of vegetation, and hence pollen assemblages. Vegetation patterns, and the distribution of individual species, are controlled by many factors including seasonal temperature and precipitation regimes, disturbance (including wildfires), and human activities. Pollen-based palaeoclimate methods have long exploited the multivariate nature of pollen assemblages to reconstruct different aspects of climate (see e.g. the discussion in Bartlein et al., 2011). The CCA shows that there is sufficient information in the pollen assemblages to assess the independent contribution of fire to vegetation assemblages. The overall relationship between pollen and charcoal-derived burnt area is reasonably strong (R^2 = 0.47), reflecting the importance of vegetation properties (gross primary production and non-tree cover) in driving the occurrence of fire – as seen in the GLM analysis of satellite-derived modern burnt area patterns. The overwhelming importance of vegetation properties in influencing modern fire occurrence is consistent with results from global analyses (e.g. Moritz et al., 2012; Pausas and Ribeiro, 2013; Bistinas et al., 2014; Forkel et al., 2019b). Nevertheless, the GLM analysis shows that climate factors, in particular the occurrence of dry intervals, are important controls on modern fire patterns in Iberia. Again, this is consistent with global analyses of the modern drivers of fire occurrence.

How successful was the Iberian test case? Not very successful, as far as I can see. The authors report that pollen data predict charcoal abundances through time "relatively well (R^2 = 0.47)".

The sentence in the abstract led to a misunderstanding here. The reported R^2 value is for the relationship between the pollen and burnt area, not the relationship between pollen and charcoal abundance through time. We will revise this sentence in the abstract to:

The pollen data predict charcoal-derived burnt area relatively well (R^2 = 0.47) and the changes in reconstructed burnt area are synchronous with known climate changes through the Holocene.

However, as Figure 4 in the original version of this manuscript shows (Fig. 5 in reply to reviewer comments) this is almost entirely due to a long-term trend during the Holocene towards increased burning. Centennial or millennial scale peaks and troughs in this graph (no longer included in the paper, nor is the helpful flow chart of methodology – why?) are mostly mis-aligned, a point made already by reviewer 1.

We have not been asked to upload a revised manuscript at this point. However, we propose to include both this Figure (and the flow-chart) in a revised manuscript. As we pointed out in our response to Reviewer 1, the choice of the loess smoothing span has an impact on the shape of the curve. The original span was chosen to emphasise the long-term trends. Furthermore, as we pointed out in our response to Reviewer 1, the number of records included in the two time series was different. The revised figure (now Figure 5) uses only data that are in common between the two data sets, and uses a smaller span for the loess *smoothing to more realistically represent shorter-term variations. This revised figure shows better congruence between the placing of peaks.*

In their reply to her/him, the authors also report the results of CCA which shows that burnt area alone explains only 1% of the variability, while other factors explain a much higher share. My guess is that much of this is due to the fact that most of the 29 sites with coupled pollen-charcoal analyses come from two relatively small mountain areas of central Spain (see Figure 1 map in the current manuscript version), so that other regions and biotypes are under-represented in the "training set".

Indeed, burnt area alone only explains a small proportion of the variability in the pollen assemblages. Nevertheless, this is variability that is not explained by other factors (i.e. it is independent information). Given that vegetation assemblages are controlled by many factors, we do not expect burnt area to explain a high proportion of the variability. Please see our responses above.

Although there are several sites from the central mountains of Spain, more than half the records are from outside these regions (see Figure 1 in original manuscript). We agree that the calibration data set is relatively small and it would indeed be worthwhile to test the relationships derived for Iberia over a wider area in order to explore this further.